

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 15:44:13 ; Search time 167.964 Seconds  
(without alignments)  
7348.085 Million cell updates/sec

Title: US-09-714-865-1

Perfect score: 2224

Sequence: 1 acctgaagtcaccatg9999.....aagctctggtttgtgatgca 2224

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS-COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	493	22.2	504	4	US-09-621-976-2923
2	389.8	17.5	420	4	US-09-833-381-1475
3	386.4	17.4	2451	4	US-09-976-594-786
4	383.2	17.2	2319	3	US-09-058-489-90
5	383.2	17.2	4416	3	US-09-058-489-17
6	369.4	16.6	3408	3	US-09-058-489-14
7	369.4	16.6	5322	3	US-09-058-489-13
8	158.2	7.1	2365	4	US-09-183-706-42
9	158.2	7.1	2365	4	US-09-567-995-42
10	129.4	5.8	1191	4	US-09-328-352-509
11	129	5.8	3760	4	US-09-576-594-213
12	127.6	5.7	1410	4	US-09-543-681A-1401
13	125.6	5.6	1830121	4	US-09-557-884-1
14	125.6	5.6	1830121	4	US-09-557-884-1
15	123.6	5.6	1254	5	PCT-US96-05320A-894
16	123.6	5.6	1830121	4	US-09-643-990A-1
17	123.6	5.6	1830121	4	US-09-643-990A-1
18	116.8	5.3	3230	4	US-08-961-527-203
19	115.2	5.2	3347	4	US-09-702-705-318
20	115.2	5.2	3347	4	US-09-726-457-318
21	115.2	5.2	3347	4	US-09-614-124B-318
22	115.2	5.2	3347	4	US-09-671-325-318
23	115.2	5.2	3347	4	US-09-589-184-318
24	115.2	5.2	3825	3	US-09-208-742-3
25	113.4	5.1	1317	4	US-09-543-681A-195
26	110.4	5.0	45613	4	US-09-596-002-22
27	110	4.9	1941	4	US-09-328-352-1891

#### ALIGNMENTS

##### RESULT 1

US-09-621-976-2923

; Sequence 2923, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 2923

; LENGTH: 504

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 224..472

US-09-621-976-2923

Query Match 22.2%; Score 493; DB 4; Length 504;

Best Local Similarity 99.8%; Pred. No. 8,9e-134; Mismatches 0; Indels 1; Gaps 1;

Matches 504; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 775 TGAGGACTCCATCTTTGCACATTATCAGACAGGCAATAAATCGACAAATACGACACTAT 834

Db 1 TGAGGACTCCATCTTTGCACATTATCAGACAGGCAATAAATCGACAAATACGACACTAT 60

QY 835 TCTTGTGGAGTGTCTGGACATGATGCACACAGCAATCTGACTTTTGAAGAGCTAA 894

Db 61 TCTTGTGGAGTGTCTGGACATGATGCACACAGCAATCTGACTTTTGAAGAGCTAA 120

QY 895 TCTTGTGGAGTGTCTGGACATGATGCACACAGCAATCTGACTTTTGAAGAGCTAA 954

Db 121 TCTTGTGGAGTGTCTGGACATGATGCACACAGCAATCTGACTTTTGAAGAGCTAA 180

QY 955 GCAAAATACAGTATTCCTATCATCTTGCAGGAGGAGATTTGATGCTTGCTCAAAC 1014

Db 181 GCAAAATACAGTATTCCTATCATCTTGCAGGAGGAGATTTGATGCTTGCTCAAAC 240

QY 1015 AGGGTCTGGGAGACTGGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGG 1074

Db 241 AGGGTCTGGGAGACTGGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGG 300

QY 1075 AATAACTCCAGTCCGTTTTAAAGAGTTGCGAGAACCGAGTGATTTATTTAGCACCAAC 1134

Db 301 AATAACTCCAGTCCGTTTTAAAGAGTTGCGAGAACCGAGTGATTTATTTAGCACCAAC 360

Sequence 1384, App

Sequence 1384, App

Sequence 1384, App

Sequence 1384, App

Sequence 19, Appl

Sequence 788, App

Sequence 1425, App

Sequence 1, Appli

Sequence 3545, App

Sequence 816, App

Sequence 29, Appl

Sequence 3206, App

Sequence 126, App

Sequence 126, App

Sequence 126, App

Sequence 126, App

Sequence 126, App

Sequence 126, App



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 15:44:13 ; Search time 167.964 Seconds  
(without alignments)  
7348.085 Million cell updates/sec

Title: US-09-714-865-1  
Perfect score: 224  
Sequence: 1 acttgaagtcaccatggggg.....aagtctgtggtttgatgca 2224

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	493	22.2	504	4	US-09-621-976-2923
2	389.8	17.5	420	4	US-09-833-381-1475
3	386.4	17.4	2451	4	US-09-976-594-786
4	383.2	17.2	2319	3	US-09-058-489-90
5	383.2	17.2	4416	3	US-09-058-489-17
6	369.4	16.6	3408	3	US-09-058-489-14
7	369.4	16.6	5322	3	US-09-058-489-13
8	158.2	7.1	2365	3	US-09-183-708-42
9	129.4	5.8	1191	4	US-09-567-995-42
10	129.4	5.8	1191	4	US-09-328-352-509
11	129.4	5.8	3760	4	US-09-976-594-213
12	127.6	5.7	1410	4	US-09-543-681A-1401
13	125.6	5.6	1830121	4	US-09-557-884-1
14	125.6	5.6	1830121	4	US-09-643-980A-1
15	123.6	5.6	1254	5	PCI-US96-05320A-894
16	123.6	5.6	1830121	4	US-09-557-884-1
17	123.6	5.6	1830121	4	US-09-643-980A-1
18	116.8	5.2	3320	4	US-08-961-527-203
19	115.2	5.2	3347	4	US-09-702-705-318
20	115.2	5.2	3347	4	US-09-736-457-318
21	115.2	5.2	3347	4	US-09-614-124B-318
22	115.2	5.2	3347	4	US-09-671-325-318
23	115.2	5.2	3347	4	US-09-589-184-318
24	115.2	5.2	3825	3	US-09-208-742-3
25	113.4	5.1	1317	4	US-09-543-681A-195
26	110.4	5.0	45613	4	US-09-596-002-22
27	110	4.9	1941	4	US-09-328-352-1891

ALIGNMENTS

RESULT 1

US-09-621-976-2923

; Sequence 2923, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 2923

; LENGTH: 504

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 224..472

US-09-621-976-2923

Query Match

Best Local Similarity 22.2%; Score 493; DB 4; Length 504;

Matches 504; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 775 TGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTAT 834

Db 1 TGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTAT 60

Qy 835 TCTTGTGGAAGTCTCTGGACATGATGCACACAGCAATTCGACATTTGAAGAAGCTAA 894

Db 61 TCTTGTGGAAGTCTCTGGACATGATGCACACAGCAATTCGACATTTGAAGAAGCTAA 120

Qy 895 TCTCTGTGCAGACACTGAATAACCAACTTGTCTAAAGCTGGTTTACTAAGCTTACTCTGT 954

Db 121 TCTCTGTGCAGACACTGAATAACCAACTTGTCTAAAGCTGGTTTACTAAGCTTACTCTGT 180

Qy 955 GCAAAATACAGTATTCCTATCATCTTGCAGACAGATTTGCATGGCTTGTCTCAAC 1014

Db 181 GCAAAATACAGTATTCCTATCATCTTGCAGACAGATTTGCATGGCTTGTCTCAAC 240

Qy 1015 AGGCTCTGGAAGACTGGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGG 1074

Db 241 AGGCTCTGGAAGACTGGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGG 300

Qy 1075 AATAACTGCCAGTCGTTTTAAAGAGTTGAGGAGCCAGAGTGTATTATTGTAGCACAAC 1134

Db 301 AATAACTGCCAGTCGTTTTAAAGAGTTGAGGAGCCAGAGTGTATTATTGTAGCACAAC 360

1135 TCGAGATTGCTCAACAGATTATTTTGGAGCCAGAAATTTCTTTGGGACTGTGT 1194  
Db 361 TCGAGATTGCTCAACAGATTATTTTGGAGCCAGAAATTTCTTTGGGACTGTGT 420  
Qy 1195 AAGAGCTGTTGTTATATATATGAGGAGCCAGCTGGGACATCAATTCGACAAATAGTACA 1254  
Db 421 AAGAGCTGTTGTTATATATATGAGGAGCCAGCTGGGACATCAATTCGACAAATAGTACA 479  
Qy 1255 AGGCTGTAATATATATGCTACT 1279  
Db 480 AGGCTGTAATATATATGCTACT 504

## RESULT 2

US-09-833-381-1475  
; Sequence 1475, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1475  
; LENGTH: 420  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1475

Query Match 17.5%; Score 389.8; DB 4; Length 420;  
Best Local Similarity 99.5%; Pred. No. 9.9e-104;  
Matches 391; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 818 GACAAATAGCACACTATTTCTGTGGAAGTGTCTGCACATGATGCACACAGCAATTCGT 877  
Db 28 GTCAATACGACACTATTTCTGTGGAAGTGTCTGCACATGATGCACACAGCAATTCGT 87  
Qy 878 ACTTTGAGAGCTAACTCTCTGCAGACACTGAATAACACATTCGTAAGCTGGTAT 937  
Db 88 ACTTTGAGAGCTAACTCTCTGCAGACACTGAATAACACATTCGTAAGCTGGTAT 147  
Qy 936 ACTAAGCTTACTCTCTGCAGAAATACAGTATTCCTATCATCTGTCAGGACGAGATTG 997  
Db 148 ACTAAGCTTACTCTCTGCAGAAATACAGTATTCCTATCATCTGTCAGGACGAGATTG 207  
Qy 998 ATGGCTTTGTCTCAACAGGGTCTGGGAAGCTGGGGCTTTTCTCTACCAATTTGGCT 1057  
Db 208 ATGGCTTTGTCTCAACAGGGTCTGGGAAGCTGGGGCTTTTCTCTACCAATTTGGCT 267  
Qy 1058 CATATGATCATGATGGAATACTGTCAGTCTGTTTAAAGAGTTCAGGACACAGAGTGT 1117  
Db 268 CATATGATCATGATGGAATACTGTCAGTCTGTTTAAAGAGTTCAGGACACAGAGTGT 327  
Qy 1118 ATTATTGTAGCCAACTCGAGAATTGGTCAACAGATTATTTTGGAGCCAGAAAATTT 1177  
Db 328 ATTATTGTAGCCAACTCGAGAATTGGTCAACAGATTATTTTGGAGCCAGAAAATTT 387  
Qy 1178 TCTTTGGAGCTTGTTAAGAGCTGTTGTTATA 1210  
Db 388 TCTTTGGAGCTTGTTAAGAGCTGTTGTTATA 420

## RESULT 3

US-09-976-594-786  
; Sequence 786, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael

APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROI  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PBR L Program  
; SEQ ID NO 786  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; FEATURE:  
; OTHER INFORMATION: Incyte ID No. 6673549 035282CB1  
US-09-976-594-786

Query Match 17.4%; Score 386.4; DB 4; Length 2451;  
Best Local Similarity 58.7%; Pred. No. 2.7e-102;  
Matches 763; Conservative 0; Mismatches 501; Indels 36; Gaps 4;  
Qy 801 AGACAGGCATAAATTCGACAAATACGACACTATTTCTTGTGGAAGTGTCTGACATGATG 860  
Db 503 ACACGGGATTAACTTTGAGAAATATGATATATACAGTAGAGGCAACCGGACGTAATC 562  
Qy 861 CACACACGACATTCCTGACTTTTGNAGAGCTATCTCTGCAGACACTGATATACACACA 920  
Db 563 GTCCCTCCACATATTGAGAAATTTTACGATATTGACATGGGAGAAATTTATCATGGGGAACA 622  
Qy 921 TTGCTAAAGCTGGTTATATAAGCTTACTCTCTGTGCAAAAAATACAGTATTCCTATCATAC 980  
Db 623 TTGAACCTTACTGCTATATCTGCTCTACTCCAGTGCAAAAACATGCCATTCCTATTATTA 682  
Qy 981 TTGACAGGACGATTTGATGGCTTGTCTCAACAGGCTCTGGGAGAGCTGGGGCTTTTC 1040  
Db 683 AGGGAAAAAGAGACTTAATGCTTGTGCCAAACAGGATCTGGGAAAACTGCAGCATTTTC 742  
Qy 1041 TCTACTCAATTTTGGCTCA-----TATGATGCATGATG 1073  
Db 743 TTTTACCCTACTGACTCAGATATATACAGATGGTCCAGGAGAGCTTTTGAAGGCTGTGA 802  
Qy 1074 GAATTAATCCAGTCTGTTTTAAAGAGTTGCAGGACACAGAGTGATTTATTGTAGCACCAA 1133  
Db 803 AGGTAAATGGAGGATGATGGCCCGCCAAACAATATCCAAATATCTTGGTTTTAGCCCCAA 862  
Qy 1134 CTCGAAATGGTCAACACAGATTATTTTGGAGCCAGAAAAATTTCTTTTGGGACTTGTG 1193  
Db 863 CAGAGAAATGGCTGTACAGATCTATGAGGAGCCAGAAAAATTTTCTTACCATCTAGAG 922  
Qy 1194 TAAGAGCTGTTGTTATATATATGAGGGAACCCAGCTGGGACATTCATTTGACAAATAGTAC 1253  
Db 923 TTGCTCTTGTGTAGTTTATGGTGTGTGATATTGGTCCAGCAGATTTCGGGACTTAGAAC 982  
Qy 1254 AAGGCTGTAATATATATGCTGCTACTCTGGAAGACTGATGGATATCATAGCAAGAA 1313  
Db 983 GTGGATGCCACTTGTGTAGTACCTCCAGGACGCTAGTGGATATGATGGAAAGAGAA 1042  
Qy 1314 AGATTGCTCTCAACAGATCAATATCTTAGTTTGGATGAAGCTGATGCTATGTTGGATA 1373  
Db 1043 AGATTGGATTAGACTTCTGCAAGTACTTAGTCTTGGATGAAGCTGATAGTACTGTGATA 1102  
Qy 1374 TGGTTTTGGTCCAGAAATGAAGATTTAATTTCTTGGCCAGGAATGCCATCAAGGAAC 1433  
Db 1103 TGGGATTTGAACCTCAGATACGCTGATAGTTGAACAAGATACTATGCCACCAAGGGCG 1162  
Qy 1434 AGCCCAAAACCTTATGTTTCAGTGGCAACTTTTCCAGAGGAAATTTCAAAGTTGGCTGCAG 1493  
Db 1163 TTGCTCACACCATGATGTTTAGTGTACTTTTCTTAAGGAATACAGATGCTTGTCTCGTG 1222  
Qy 1494 AGTTTTAAAGTCAAAATATCTGTTGTTGCTGTTGGACAGAGTGGTGAGCATGTAGAG 1553  
Db 1223 ACTTTTGGATGAA---TATATCTTTTGGCTGTAGGACAGTAGGCTTACTCTCTGAGA 1279

```
QY 1554 ATGTTACGACGAGCGTTCTCCAAAGTGGCCAGTTCTCAAAAAGAGAAAGCTCGTTGAAA 1613
Db 1280 AATCAACACAGAGAGTAGTTGGTGGAGAGACTTAGATTAACGGTCAATTTCTACTGGACA 1339
QY 1614 TTCTGCGAAA CATAGGG---GATGAAAGAACTATAGTCTTTGTTGTTGAAACTTAAGAAAG 1670
Db 1340 TTTTAGTGTCACACAGGAGTAGTCACTTACTTTAGTGTGTTGGAGACCCAAAAGGGAG 1399
QY 1671 CAGATTTTACTGCAACTTTCTTTGTCAGAAAGAAATATCAACTACAGTATCCATGGTG 1730
Db 1400 CAGATTCCTCGAGGATTTCTTATACCATGAAGATATGCTTGTACTAGTATTCATGGAG 1459
QY 1731 ATCGGAAACAGAGAGCGGAGCAAGCTTTGGAGATTTTCGTTTGGAAAGTCCCGAG 1790
Db 1460 ACCGCTCACAGAGATCGAGAGGAGGCCCTTCCAGTTTCGCTCAGGAAAAAGCCCAA 1519
QY 1791 TTCTGTTGCTACTTCAGTAGTCCGAGAGGCTGGATATTAAGAAATGTCACATCTTA 1850
Db 1520 TTCTAGTGGCTACAGCTGTGGCAGCAGAGACTAGACATTTCAAAATGTGAGACATGTTA 1579
QY 1851 TCAATTTTGTATCTTCTTCTACCATTTGATGATATGTTTCATCGAATTTGGGCGTACTGGTC 1910
Db 1580 TCAATTTTGTATTTGCAAGTATATTTGAAGATATGTCATCGTATTTGGCCGTACAGGAC 1639
QY 1911 GTTGGGAAATCTGGCAGAGCAATTTCTTTTGTATCTTGAATCGGATACCATTTAG 1970
Db 1640 GTGTAGGAAACCTGGGCGCTTGGCCCTCATTTCTTTAA---TGAAGAAATATGAATATTA 1696
QY 1971 CACAGCTCTAGTAAAGTATTGACAGATGCTCAACAGGATGTTCTCGATGTTGGAAG 2030
Db 1697 CAAAGATTTGTTGGATCTTCTTGTAGAGCTTAACAGAGTGGCTTCTTGGTTGAAA 1756
QY 2031 AAATGTCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2070
Db 1757 ATATGGCTTATGAACACCACTACAGGGGTGGCAGTCTGGT 1796

RESULT 4
US-09-058-489-90
; Sequence 90, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WH197-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-90

Query Match 17.2%; Score 383.2; DB 3; Length 2319;
Best Local Similarity 58.5%; Pred. No. 2.2e-101;
Matches 761; Conservative 0; Mismatches 503; Indels 36; Gaps 4;

QY 801 AGACAGCGATAACTCGNCAATATACGACACTATTCTTGTGGAGTGTCTGGACATGATG 860
Db 529 ACACGGGGATTAACCTTTGAGAAATATGATATACCATAGTAGAGCAACCGGCGTAGTACT 588
QY 861 CACCAACCGCAATTTCTGACTTTTGAAGAGCTTAATCTCTGTGACACACTGAATACACA 920
Db 589 GTCTCTCCATATTCGAATTTTACGATATTTGACATGGAGAAATTTATCATGGGGAACA 648
```

```
Db 1723 CAAAGGATTTGGTGGATCTCTTTGTAGAACTAAACAAGAGTGCCTTCTTGGTTGGAAA 1782
Qy 2031 AATTGGCTTTAGTACATACATCTCCCTGGCTTCAGTGTAG 2070
Db 1783 ATATGGCTTATGAACACCACTACAAAGGTGGCAGTCTGG 1822

RESULT 5
US-09-058-489-17
; Sequence 17, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 4416
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-17

Query Match 17,24; Score 383.2; DB 3; Length 4416;
Best Local Similarity 58.5%; Pred. No. 3.2e-101; Indels 36; Gaps 4;
Matches 761; Conservative 0; Mismatches 503;

Qy 801 AGACAGGCATAACTTCGACAAATACGACACTATTTCTGGAAGTGTCTGGACATGATG 860
Db 529 ACACGGGATTAACCTTTGAGAAATATCATGATATACAGTAGAGGCAACCGCGAGTAAC 588
Qy 861 CACACACGCAATTTCTGACTTTTGAAGAGCTATCTCTGACACACTCAATAACAACA 920
Db 589 GTCCTCCACATATTGAGAAATTTAGCGATATTGACGGAGAAATATCATGGGNAACA 648
Qy 921 TTGCTAAAGCTGGTTATCTAAGCTTACTCTGTGCAAAATACAGTATTCCTATCATAC 980
Db 649 TTGAACCTTACCTGCTATCTCTCTACTCTGAGTGCACAAACATGCCATTCCTATTATTA 708
Qy 981 TTGCAGGACGAGATTTGATGCTTGTCTCAACAGGGTCTGGGAAGACTCGCGCTTTTC 1040
Db 709 AGGGAAAAGAGACTTATGCTGTGCTTGTGCCCCAAACAGGATCTGGGAAACCTGCAGCATTC 768
Qy 1041 TCCTACCAATTTTGGCTCA-----TATGATGCATGATG 1073
Db 769 TTTTACCCNTACTGAGTCAGATATATACAGATGCTCCAGGAGAGCTTTGAAGCTGTGA 828
Qy 1074 GAATAACTGCCAGTGGTTTAAAGAGTTGCGAGGAAACGAGTGTATTATTGTAGCACCAA 1133
Db 829 AGGAAATGGAAGGTATGGGCGCGCAACAAATATCCAATATCTCTGGTTTGAACCCCAA 898
Qy 1134 CTGAGGATTTGTCACACAGATTTATTTGGGAAGCCAGAAATTTTCTTTGGGACTTTGTG 1193
Db 889 CAAGAGAAATGGCTGTACAGATCTATGAGGAGCCAGAAATTTTCTTACCAGTCTAGAG 948
Qy 1194 TAAGAGCTGTTGTTATATATATGGGGAACCCAGCTGGGACATTCATATTCGACAAATAGTAC 1253
Db 949 TTCGTCCTTGTGTAGTTTATGCTGTGCTGATATTTGGTTCAGCAGATTCGGGACTTAGAAC 1008
Qy 1254 AAGGCTGTAATATATATGCTACTCTCTGGAAGACTGATGATATCATAGGCAAGAAA 1313
Db 1009 GTGGATGCCACTTGTGTAGTCCCTCCAGGACGCTGTAGTGTATGATGAAGAGGAA 1068
Qy 1314 AGATTGGTCTCAACAGATCAATATCTAGTTTGGATGAAGCTGATGCGATGTTGATA 1373
```

```
Db 1069 AGATTGGATTAGACTTCTGCAAGTACTTAGTGTGGATGAAGCTGATAGGATGCTGGATA 1128
Qy 1374 TGGTTTTTGGTCCAGAAATGAAGAAGTAAATTTCTCCGCCAGGAATGCCATCAAGGAAC 1433
Db 1129 TGGGATTTGAACCTCAGATAGCTGCTATAGTTGAACAAGATCTATATCCCAAGGGGCG 1488
Qy 1434 AGCGCCAAACCCCTTATGTTCACTGTCGAATTTTCCAGAGGAAATTCAAAGGTTCGCTCAG 1493
Db 1189 TTGTCACACCATGATGATTTAGTGTCTACTTTTCTTAAGGAAATACAGATCTTGTCTG 1248
Qy 1494 AGTTTTTAAAGTCAAAATATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1553
Db 1249 ACTTTTGGATGAA---TATATCTTTTGGCTGTAGGCAGAGTAGGCTCTACTCTCTGAGA 1305
Qy 1554 ATGTTTCAGCAGACCGTTCTCCAACTTGGCAGATTTCTCAAAAGAGAGAAAAGTCTCGTTGAAA 1613
Db 1306 ACATCACACAGAAAGTAGTTTGGGTGGAAGACTTAGATTAACGGTCAATTTCTTACTGACA 1365
Qy 1614 TTCTGCGAAACATAGGG---GATGAAAGAACTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1670
Db 1366 TTTTAGGTGCAACAGGAGTGTATCTACTTTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 1425
Qy 1671 CAGATTTTACTGCAACTTTTCTTCTCAAGAAAAAATATCAACTCAAGTATCCATCGTG 1730
Db 1426 CAGATTCCTCGAGGATTTCTTATACCATGAAGGATATGCTTTGTACTAGTATTCATGAG 1485
Qy 1731 ATCGGGAACAGAGAGCGGAGCAAGCTCTTGGAGATTTTTCGCTTTTGGAAAAGTCCCGAG 1790
Db 1486 ACCGTCACAGAGATCGAGGAGGCCCTTCCACAGTTTCGCTCAGGAAAAAGCCCAA 1545
Qy 1791 TTCTTGTGCTACTTCAGTAGTGTGTCAGAGGCTGGATATTTGAAAATGTGCAACATGTTA 1850
Db 1546 TTCTAGTGGCTACAGCTGTGCGAGCAGCGAGACTAGACATTTCAATATGTGAGACATGTTA 1605
Qy 1851 TCAATTTTGTGCTTCTTCTTACCATGATGATATGTTTCATCGAATTTGGGGTACTGCTGC 1910
Db 1606 TCAATTTTGTGCTTCTTCTTACCATGATGATATTTGAAGAAATGTCATCGTATTTGGCCGTACAGAC 1665
Qy 1911 GTTGTGGGAATCTGCGAGAGCAATTTCTCTTTTGTGATCTTTGAAATGCGAATAACCATTTAG 1970
Db 1666 GTGTAGGAAAACCTGGGCTTGCACCTCATTTCTTAA---TGAATAAATAATGAATATTA 1722
Qy 1971 CACAGCTCTAGTAAAGTATTGACAGATGCTCAACAGGATGTTCTCTGATGTTGGAAG 2030
Db 1723 CAAGGATTTGTTGGATCTTCTTGTAGAAAGCTTAAACAAGAGTGCCTTCTTGTGTTGAAA 1782
Qy 2031 AAATTGCTTTAGTACATACATTTCTGCTTCAGTGTGTTAG 2070
Db 1783 ATATGCTTATGAACACCACTACAGAGGTGGCAGTCTGTG 1822

RESULT 6
US-09-058-489-14
; Sequence 14, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3408
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-14
```

Qy	1776	TTGGAAGTGC	CGAGTTC	TTGTTGCT	CTACTCT	AGTAGT	CGCAGAGGGCT	GGATAT	TGAA	1833				
Db	2322	CAGGAAAAAG	CCCAAT	TTT	TAGTGGCT	TACAGCAGT	AGCAGAC	TGGACAT	TTTCAA	2381				
Qy	1836	ATGTGCAAC	ATGTTAT	CAATTT	TGATCT	TCTT	TACCAAT	TGATGAAT	TGTTTCAT	TCGAA	1895			
Db	2382	ATGTGAAAC	ATGTTAT	CAATTT	TGATCT	TCTT	TACCAAT	TGATGAAT	TGTTTCAT	TCGAA	2441			
Qy	1896	TTGGGGCT	ACTGCT	GTGTGG	GAAT	PACT	GGCAGAGCA	ATTCCT	TTTTTT	TGATCT	TGAAT	1955		
Db	2442	TTGGT	CGTACGG	ACGCT	GTAGG	AAAA	CCCTT	GGCCTG	GCACAC	CTCAT	TTCTTTAA	2498		
Qy	1956	CGGAT	AACCAATTT	TAGCAC	AGCCT	CTAGT	AAAA	AGTATT	GACAGAT	GTCT	CAACAGGAT	GTTC	2015	
Db	2499	GGAA	CATAA	TAT	TACT	AGAT	TTGTT	TGATCT	TCTT	TGNAAGT	ATAACA	AGAGTGC	2558	
Qy	2016	CTGCAT	GGTTGG	AGAAAT	TGCC	TTT	TAGTAC	ATACAT	TCTCGC	TTT	CAGTGG	TAG	2070	
Db	2559	CGT	TTGGTT	TAGAAA	CAT	GGCT	TAT	GAA	CACCACT	TACAA	GGGTAG	CAGT	CGTGG	2613
RESULT 7														
US-09-058-489-13														
; Sequence 13, Application US/09058489														
; Patent No. 6103886														
; GENERAL INFORMATION:														
; APPLICANT: Whitehead Institute for Biomedical Research														
; APPLICANT: Lahn, Bruce														
; APPLICANT: Page, David														
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of														
; FILE REFERENCE: WH197-08pA														
; CURRENT APPLICATION NUMBER: US/09/058,489														
; CURRENT FILING DATE: 1998-04-10														
; EARLIER APPLICATION NUMBER: 60/041,877														
; EARLIER FILING DATE: 1997-04-11														
; NUMBER OF SEQ ID NOS: 91														
; SOFTWARE: FastSeq for Windows Version 3.0														
; SEQ ID NO 13														
; LENGTH: 5322														
; TYPE: DNA														
; ORGANISM: Human														
US-09-058-489-13														
Query Match 16.6%; Score 369.4; DB 3; Length 5322;														

Query Match	16.6%	Score 369.4	DB 3	Length 3408
Best Local Similarity	57.6%	Pred. No. 2.9e-571	Indels 36	Gaps 4
Matches 758	Conservative 0	Mismatches 521		
QY	786	TCTTTGCACATTATCAGACAGGCATAACTTCGACAAATACGACACTATTCTTGTGGAAG	845	
DB	1305	TCTTTTCGAGGCAACA CTGGGATTAAATTTTGAGAAATACGATGACATTCCTCAGTTGAGG	1364	
QY	846	TGTCCTGGACATGATGCACACCGACAACTCTGACTTTTGAAGAGCTAATCTCTGTCAGA	905	
DB	1365	CAACAGGCAACAACCTCTCTCCACATATTGAAAGTTTCAGTGATGTTGAGATGGGAGAA	1424	
QY	906	CAC TGAATAACCAACATTGCTTAAAGCTGGTTATACTAAAGCTTACTCTCTGTCGCAAAAATCA	965	
DB	1425	TTATCATGGMAACATTGAGCTTACTCGTATATACTCGCCCAACTCCAGTGCCTAAAGCATG	1484	
QY	966	GTATTCCTATCATCTTGCAGGACGAGATTGAGTGTCTGCTCAAAACAGGCTCTGGGA	1025	
DB	1485	CTATTCCTATTAACAAGAGAAAGAGACTTGAATGCTTGTGCCCAACAGGCTCTGGAA	1544	
QY	1026	AGACTCGGGCTTTCTCTCAACCAATTTTGGCTCATATGATGCAATGAGTGG-----	1074	
DB	1545	AAACTCGAGCATTTCTGTTGGCCCATCTTGAGTCAGATTTATTTCAGATGGTCCAGGCGAGG	1604	
QY	1075	-----AATAACTGCCAGTCGTTTTAAAGAGTTTGCAGGACACAGAGTGTA	1118	
DB	1605	CTTTGAGGGCCATGGAAGAAATGGAAGGTATGGCGCGCCGCAACAACTACCCATCTCTCT	1664	
QY	1119	TTATTGTAGCACCACACTCGAGAAATGGTCAACAGATTTATTTTGAAGCGCAGAAAAATTTT	1178	
DB	1665	TGCTATTAGCACCACAGAGAGTTGGCAGTACAGATCTACGAAAGAGCGCAGAAAAATTTT	1724	
QY	1179	CTTTTGGGACTTGTGAAGAGCTGTTGTTATATATGSGGAACCCAGCTGGGACATTCAA	1238	
DB	1725	CATACCGATCTAGAGTTGTCCTTGGTGTTTATGGTGTGCCGATATTTGGTCAAGCA	1784	
QY	1239	TTCCGACAAATPAGTACAAGGCTGPAATATATATATGTGCTACTCTCTGGAAGACTGATGATA	1298	
DB	1785	TTCCGAGACTTGGAAACGTGGATGCCATTTGTTAGTAGCCAATCCAGGACGCTCTAGTGGATA	1844	
QY	1299	TCATAGGCAAGAAAGATTTGGTCTCAACAGAGATCAAACTACTTTAGTTTTCGATGAAGCTG	1358	
DB	1845	TGATGGAAGAGGAAGATTTGGATTAGACTTTTTCGAAATACTTTGGTTAGATGAAGCTG	1904	
QY	1359	ATCCGATGTTTGGATATGGGTTTTTGGTCCGAAATGAAGAAGTTAAATTTCTTGCCCAAGGAA	1418	
DB	1905	ATCCGATGTTTGGATATGGGTTTTGAGCCTCAGATTTCTGATAGTTCGAAACAAGATACTA	1964	
QY	1419	TGCCATCAAGGAACACGCGCAACCTTATGTTTCACTGCAATTTTTCAGAGGAAATTC	1478	
DB	1965	TGCCATCAAGGGTGTCCGCAACTATGATTTAGTGCTACTTTTCTTAAGGAATATC	2024	
QY	1479	AAAGGTTGGCTCGCAGAGTTTTTAAGGTCAAATATCTGTTTGTGTTTGGTGGCAAGTGG	1538	
DB	2025	AGATGCTGGCTCGTGATTTCTTAGATGAA---TATATCTTTCTGGCTGTAGGAAGAGTTG	2081	
QY	1539	GTGGAGCATGTAGAGATGTTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCTCAAAAGAG	1598	
DB	2082	GCTCTACTCTGAAACATCACACGAAAGTAGTTTGGTGGGAAGATCAACAACGGT	2141	
QY	1599	AAAAGCTGTTGAAATTTCTGGAACATAGG---GGATGAAGAACTATTCGCTCTTTGTTG	1655	
DB	2142	CAATTCCTGCTTGAOCCTCTAAATGCAACAGGCAAGGATTCACCTGACCTTAGTGTGTG	2201	
QY	1656	AAACTAGAAAAAGCAGATTTTACTGCACTTTTCTTGTCTCAAGAAAAAATATCAACTA	1715	
DB	2202	AGACCAAAAGGGTGCAGATTTCTGAGGATTTCTTATACCTGAGGATACCGCATGTA	2261	
QY	1716	CAAGTATCCATGGTGTATCGGGAACAGAGAGCGGGAGCAAGCTCTTTGGAGATTTTCGCT	1775	
DB	2262	CCAGCATCCATGGAGACCGTTCTCAGAGGGAATAGGAAGAGGCCCTTACACAGTCCGCT	2321	



Db	1665	TTGTTATTAGCACCAACGAGAGAGTTGGCAGTACAGATCTACGAAGAAGCCAGAAAATTTT	1724
Qy	1179	CTTTTTCGGGACTTCGTGTAAAGAGCTGTTGTTATATATGSGGGAACCCAGCTGGCACATCOAA	1238
Db	1725	CATACCAGTCTAGAGTTGCTCTTTCGCTGGTTATAGTGGTGCCGATATTGGTCAGCAGA	1784
Qy	1239	TTCCGACAAATAGTATCAAGGCTGTAATATATATGTGCTACTCTCTGGAAAGCTGATGGATA	1298
Db	1785	TTCCGAGACTTTGGAAACGTGGATGCCATTTGTTAGTAGCCACTCCAGGACGCTCTAGTGGATA	1844
Qy	1299	TCATAGGCAAGAAGAAAGATTTGGTCTCAACACAGATCAAAATACATTAGTTTGGATGAAGCTG	1358
Db	1845	TGATGTAAGAGAGAAAGATTGGATTAGACTTTTGGCAATACATTGGTGTAGATGAAGCTG	1904
Qy	1359	ATCCGACTGTTGGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTCCCAGGAA	1411
Db	1905	ATCCGATGTTGGATATGGGTTTTGAGCCTCAGATTCGTAGATAGTCGAACAAGATACTTA	1964
Qy	1419	TGCCATCAAAAGGAACAGCGCCAAACCCATTATGTTCAGTCGCACTTTTCCAGAGGAAATTC	1478
Db	1965	TGCGCTCAAAAGGGTGTCCGCGCACACTATGATGTATTAGTGTCTACTTTTCTTAAGGAATAC	2024
Qy	1479	AAAGTTGGCTGCAGAGTTTAAAGTCAAAATATCTGTTTGTGCTGTGCAAGTAGTG	1538
Db	2025	AGATGCTGGCTCGTGATTTCTTAGATGAA---TATATCTTCTTGGCTGTAGGAAGTTG	2081
Qy	1539	GTGGAGCATGTAGAGATGTTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAGAG	1598
Db	2082	GCTCTACTCTGTAANAACATCACACAGAAAGTAGTTTGGGTGGGAAGAATCAGACAAACGGT	2141
Qy	1599	AAAAGCTGTTGAAATTCGCGAAAACATAGG---GGATGAAGAACTATGCTCTTTGTTG	1655
Db	2142	CATTCTGCTTGACCTCCTAAATGCAACAGCGAAGGATTCACGTGCCTTAGTGTGTTGG	2201
Qy	1656	AAACTAAGAAAAAGCAGATTTTACTGCACACTTTTCTTCTCAAGAAAAAATCAACTA	1715
Db	2202	AGACCAAAAAGGGTGCAATTTCTCTGGAGATTTCTTATACCATGAAGNTACGCATGTA	2261
Qy	1716	CAAGTATCCATGTGTGATCGGGAACAGAGAGCGGGAGCAAGCTCTTTGGAGATTTTCGCT	1775
Db	2262	CCAGATCCTCATGGAGACCGTTCTCAGAGGGATAGAGAAAGGGCCCTTCCACAGTTCCGCT	2321
Qy	1776	TTGGAAAGTGCCAGTTCTTGTGTCTACTTTCAGTAGCTGCGCAGAGGCGTGGATATTGAAA	1835
Db	2322	CAGGAAAAAGCCCAATTTTATGCTGCTACAGCATGACGACGACAGAGACTGGACATTTCAA	2381
Qy	1836	ATGTGCAACATGTTTATCAATTTTGAATCTTCTTACCATTGTATGAATATGTTTCATCGAA	1895
Db	2382	ATGTGAACAATGTTTATCAATTTTTCAGTTGCCAAGTGATAATTGAAGAATATGTACATCGTA	2441
Qy	1896	TTGGGCGTACTGGTTCGTTGTTGGGTAATCTGGCAGAGCAATTTCTTTTGTGATCTTTGAAT	1955
Db	2442	TTGGTCTGATCGGGAACGTTGAGGAACCTTGGCCCTGGCAACCTCATCTCTTTAA---CGAGA	2498
Qy	1956	CGGATAACCAATTTAGCACAGCCTCTAGTAAAGATTTAGACAGATGCTCAACAGGATGTTTC	2015
Db	2499	GGACATATATATTACTTAAGGATTTGTTGGATCTTCTTGTGTAAGCTTAACAGAAGTGC	2558
Qy	2016	CTGCATGGTTGGAAAGAAATTTGCCTTTTAGTACATACATTCCTCGCTTCAGTGGTAG	2070
Db	2559	CGTCTTGGTTAGAAAAACATGGCTTATGAACACCACTACAAGGGTAGCAGTCGTGG	2613

## RESULT, T 8

RESOLUT 8  
US 09-183-706-42 Application US/09183706  
; Sequence 42124525  
; Nucleotide Sequence  
; GENERAL INFORMATION:  
; APPLICANT: Martelange, Valrie  
; APPLICANT: De Smet, Charles  
; APPLICANT: Boon-Falloux, Thierry  
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR

```

/ FILE REFERENCE: L0461/7054
/ CURRENT APPLICATION NUMBER: US/09/183.706
/ CURRENT FILING DATE: 1998-10-30
/ EARLIER APPLICATION NUMBER: 09/122.989
/ EARLIER FILING DATE: 1999-07-27
/ NUMBER OF SEQ ID NOS: 43
/ SEQ ID NO 42
/ LENGTH: 2365
/ TYPE: DNA
/ ORGANISM: H. sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (208)...(2151)
US-09-183-706-42

```

Query Match	7.1%	Score 158.2	DB 3	Length 2365		
Best Local Similarity	50.1%	Pred. No. 9.2e-36				
Matches 518	Conservative 0	Mismatches 48	Indels 27	Gaps 4		
Qy	915	ACAA	CATTGCTAAAGCTGGTTATAC	TAAAGCTTACTCCTGTGCAAAAAATACAGTATTCCTA	974	
Db	971	AAAA	CATTAAAGAGCAGGTTTT	CAAAAGCCAAACACTATT	CAGTCACAGGCATGGCCCA	1030
Qy	975	TCATA	CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAAACAGGGTCTCGGAAGACTGCGG	1034		
Db	1031	TTGTG	TGCAAGGAATAGATCTTTATAGAGTAGCCAGACTGGAAACAGGAAGAACAATTGT	1090		
Qy	1035	CTTTT	CTCTACCAATTTTGGCTCATATGATGCATGGAATACTGCCAGTCGTGTTTA	1094		
Db	1091	GTTAT	TATTAATGCTGGATTTATTCACTGGTCCTCAAC	-----CCAGCCTTAAAG	1141	
Qy	1095	AAGAG	TGCGAGAAACGAGAGTGATATATGTAGACCAACCTCGAGAAATGGTCACACAGA	1154		
Db	1142	GTCA	AGGAATAGAACCCGCGCATGTTAGTCTTAACTCCCACTCGGGAATTAGCACTTCAAG	1201		
Qy	1155	TTTAT	TGGAAGCCAGAAAAATTTCTTTTGGGACTTGTGTAAGAGCTGTGTATATATG	1214		
Db	1202	TAGA	AGGAAATGTTGCAAAATATTCATAT	--AAAGGGCTTCGGAGTGTGTGTATATG	1258	
Qy	1215	GGGGA	ACCAGCTGGGACATCAATTCGACAAATAGTACAGGCTGTAAATATATATATG	1274		
Db	1259	GTGG	TGGAAATAGAGATGAACAAATAGAAAGAGCTTTAAAAAAGGTGTAGATATCATAAATG	1318		
Qy	1275	CTACT	CTCGAAGACTGATGGGATATCATAGGCAAGAAAGAAATGGTCTCAACACAGATCA	1334		
Db	1319	CAACT	CCCGAAGATTTGAATGATCTGCAATGAGTAACTTCGTCATCTGAAGATATAA	1378		
Qy	1335	AATAC	TAGTTTGGATGAAGCTGATTCGGATATGGGTTTTGGTCCAGAAATGA	1394		
Db	1379	CCTAC	TCTGTTTTAGATGAAGCAGACAGAGATGTGGACATGGGATTTGAAACCCAGATAA	1438		
Qy	1395	AGAAG	TTAATTTCTTGCCAGGAATGCCATCAAGGAAACAGCGCCAAACCCCTTATGTTCA	1454		
Db	1439	TGAAG	ATTTTGT-----TAGATGTCCGCCAGATAGGACAGACAGTTATGACCA	1486		
Qy	1455	GTGCA	ACTTTTCCAGAGGAAATTCAAAGTTGGCTGCGAGTTTTT-----AAAGTCAAAT	1511		
Db	1487	GTGCT	ACATGGGCTCATTCAGTTTCATCGCTCGCAACTCTTATTGGAAGAACCAATGA	1546		
Qy	1512	ATCTG	TTTCTGCTGTCGACAGTGGGTGGAGCATGTAGACATGTCAGCAGACCGTTC	1571		
Db	1547	TTGCT	ATGTTGGTATGTCATTTGATCTAGTTGCTGTAAAGTTCAGTGAAGCAAAATATATG	1606		
Qy	1572	TCCAG	TGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATCTCGGAAACATAGGGG	1631		
Db	1607	TAAC	CACCAGGAAGAGAAATGGAGTCAATGCAAAACTTTTCTACAGAGTATGTGATCCA	1666		
Qy	1632	ATGAA	AGAACTATGGTCTTTGTTGAACTTAAGAAAAAGCAGATTTTACTGCAACTTTTC	1691		
Db	1667	CAGAC	AAAGTCATTTGCTTCGTTTCTCGAAAGAGCTGTTCGGATCACTTATCAGTGACC	1726		
Qy	1692	TTTGT	CAAGAAAAATATCAACTACAGTATCCATGGTGATCGGAAACAGAGAGACGGG	1751		





940 TAACTTACTCTGTCAGAAATACAGTATCTCTATCATATCTGAGGACGAGATTTGAT 999  
Db TCAATGACCCCAATTCAGCAAAAGGTTTTAAATATATATATAGCGGCGATGTCAT 170  
QY 1000 GGCTTGTCTCAACAGGCTCTGGAGACTCGGCTTTCTCTACCAATTTGGCTCA 1059  
Db 171 TGGGCGAGACACAGACAGGATGAGACTGCTGCTTTCTGATAGTGTAAATTAAGA 230  
QY 1060 TATGATGCATGATGGAATTAAGTCCAGTGGTCTGTTTAAAGAGTTCAGGACACGAGTGTAT 1119  
Db 231 TTTGCTCAATTAATCCGGTTCAGAGACAGCTTCCTGGT-----GACCTCGTGT 284  
QY 1120 TATTGAGACCACTCGAATTTGGTCAACAGATTTATTTGAGAGCCAGAAATTTTC 1179  
Db 285 AATCTAGCACCTACTCGTGGTGGATTCGATTCGAGATTGAAGTGTGCAAAATCTCTAC 344  
QY 1180 TTTTGGGACTGTGTAGAGCTGTGTATATATGAGGGAACCCAGCTGGGACATTCAT 1239  
Db 345 AAAATTTCTAATTTACCTAGTACGCTATTAGGTGGTGTGATTTTCGATAGCAGAA 404  
QY 1240 TCGACAAATAGTACAAAGCTGT---AATATATTATGTCTCTCTCTGGAAGACTGATGA 1296  
Db 405 GAAGCACTTGATGCTAATTTTCTGCTATTTATGTTGCCACACAGGCGGTTAATGA 464  
QY 1297 TATCATAGGCAAGAAAGATTTGGTCTCAACAGATCAATATCTTATGTTTGGATGAGC 1356  
Db 465 TTTTGTAGAACAAAGAGATTTGGGCTCGATCAAAATGAAATTTTATGATGATGAGC 524  
QY 1357 TGATCGCATGTTGGATATGGTCTGCTCGAGAAATGAAGAGTAAATTTCTTGGCCAGG 1416  
Db 525 TGACCGTTTATGATATGGGCTTTATCTCTCGGTAACGATTTGTGCGTTATTCATCA-- 582  
QY 1417 AATGCCATCAAGAGACAGCGCCAAACCTTTATGTTCAAGTCACTTTTCCAGAGAAAT 1476  
Db 583 ----CCACGTAAGCAACAGCTCAACCTTAATGTTCTCTGCAACATTTAGTTATGATG 638  
QY 1477 TCAAGGTTGGTGCAGATTTTAAAGTCAATTTATCTGTTGTTGCTGTTGGCAAGT 1536  
Db 639 CTTGAATTTGGCAAGCAGTGGTGA---TTGHAACAGTAACTGTTGAAATTAACCTGA 695  
QY 1537 GGGTGGAGCATGTAGAGATGTTTCAGCAGACCGTTCTCCAGTTGGCCAGTCTCAAAAG 1596  
Db 696 ACAAAGACCAATAATGATGTGCAACAACTGTTTACGTTGGCTTAAACAGATAAATA 755  
QY 1597 AGAAAGCTGTTGAAATTTCTGGAACATAGGGGATGAAGCACTATGTTCTTCTGTTGA 1656  
Db 756 TCGTCTTTTCAAGATATTTTACGTGAAGACCAATTTGATAAAGTTATGATCTTTGCCAA 815  
QY 1657 AACTAAGAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAGAAAAAATATCAACTAC 1716  
Db 816 TCGCGGTGATCAGTACGTCTCTTTATGACCAATTTGAAAAGATGGATATAAAGTGG 875  
QY 1717 AAGTATCCATGATCGGAAACAGAGAGACGGGACAGCTCTTTGGAGATTTTCTGCTT 1776  
Db 876 GATGCTATCTGGTGAATTTCTCAAGATTAACCTTTTAAATAATGTTAGAGCAGTTAAGCA 935  
QY 1777 TGGAAAGTGGCCAGTCTGTTGTTGCTACTTCAGTAGCTGCGAGAGGCTGGATATGAAAA 1836  
Db 936 AGGCAACATTAACATCATGATGCAACGGATGTTGCTGCTGCTGTTATTCATGACGG 995  
QY 1837 TGTCAACATGTTATCAATTTGATCTCTCTTCAACATTTGATGATATATGTTCAATGAT 1896  
Db 996 TGTATCATGATGATGATTTTCAATTTGATGATGATGATGATGATGATGATGATGAT 1055  
QY 1897 TGGGCTAGTGGTCTGTTGGAATATCTGCGAGAGCAATTTCTTTTTCATCTGATC 1956  
Db 1056 TGGTCTGATAGCTGTCAGGGGCAAGGTGTAAATTTAGTTTCTTATCTGAAGATGA 1115  
QY 1957 GGATAACCATTTAGC 1971  
Db 1116 TGCCTTCTATTTACC 1130

RESULT 11  
US-09-976-594-213  
; Sequence 213, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Buchbinder, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROID  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 213  
; LENGTH: 3760  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 2173757CB1  
US-09-976-594-213

Query Match 5.8%; Score 129; DB 4; Length 3760;  
Best Local Similarity 48.4%; Pred. No. 3.9e-27; Indels 39; Gaps 5;  
Matches 539; Conservative 0; Mismatches 535;  
QY 844 AGTGTCTGGACATGATGCACCCAGCAATTTCTGACTTTTGAAGAACTAATCTCTCTCA 903  
Db 1236 AGTTAAAGGAAAAGTTTGCCCAACCAATTAATCTCTGGTCCAGTGTGGAATTTCCAT 1295  
QY 904 GACACTGANTAAACAACTTGTCTAAGCTGTTATCTAAGCTTACTCTCTGCAAAATA 963  
Db 1296 GAAGATCTTAAATTTCCCTCAAGAGCATGGCTATGAAAGCCCAACGCCCATCAACACCA 1355  
QY 964 CAGTATTCTCTATCATCTTGCAGACGAGATTTGATGGCTTTGCTCAACACAGGCTCTGG 1023  
Db 1356 AGCTATTCTCTATATGCTGACGAGATTTGATTTGGCATTGGCCAAACAGAGTGG 1415  
QY 1024 GAAGACTGGGCTTTCTCTCAACATTTTGGCTCATATGATGATGATGATGATGATGATG 1083  
Db 1416 AAAGACCAATCTCTCTCTGTTGCCCATGTTTAGACACATCATGATCA----- 1463  
QY 1084 CAGTCGTTTTAAAGAGTTGACAGGAAACAGAGTGTATTATTGTAGCAACCACTCGAGAA 1143  
Db 1464 GAGTCAATTAGAGAGAGAGAGGCGCAATAGCTGTCTCATGACTCTCACTCGAGACT 1523  
QY 1144 GGTCAACAGATTTATTTGGAGCCAGAAAATTTTCTTTTGGGACTTTGTGTAAGAGCTGT 1203  
Db 1524 GGCCTTACAGATTACTAAAGAGTGTAAAGAGTTTTTCCAAAGACTTTTGGGACTTTAGAGTGT 1583  
QY 1204 TGTATATATGGGNAACCCAGCTGGGACATTTCAATTCGACAAATAGTACAGGCTGTAA 1263  
Db 1584 CTGTGTTTACGAGAGAACAGGAATCAGTGAACAGATTGCTGAGCTGAAAGAGTGTCTGA 1643  
QY 1264 TATATTATGTCTACTCTCTGGAAGACTGATGATATCATAGGCAAGAAAAGATTTGCT 1323  
Db 1644 AATTATTCTTTGCACACCTGGTTCGAAATGATTGACATGTTAGCCGCTAACAGTGTGCGGT 1703  
QY 1324 CAAACAGATC-----AAATCTAGTTTGGATGAGCTGATCCATGTTGGATAT 1374  
Db 1704 CACAAATCTTGAAGAGTGCATATGTTTGTATTAGATGAAGACAGAGATTTTGGAT 1763  
QY 1375 GGGTTTTCGTCAGAAATGAAGAGTAAATTTCTTTCGCCAGGAATGCCATCAAGAGCA 1434  
Db 1764 GGGTTTGAACCCAGGTCATGCCATCGTGGTAATTTCTGT-----CCTGA 1811  
QY 1435 GGGCCAAACCCCTTATGTTCAAGTGCAATTTTCCAGAGGAATTTCAAGGTTGGTGCAGA 1494  
Db 1812 TCGACAGACGGTTATGTTTTCAGCTACTTTCCCGAGAGCTATGGAGGCTTTGGCTCGAG 1871  
QY 1495 GTTTTAAAGTCAATTTATCTGTTGTTGCTGTTTGGCAAGAGTGGTGGAGCATGTAGAGA 1554

Db 1872 GATCCTCA---GTAACCTATTGAAGTACAAAGTTGGAGCGAGGAGTGGTTTCTCAGA 1928  
Qy 1555 TGTTGAGCAGACCGTTCTCCAGTTGGCCAGTTCTCAAAAAGAGAAAGCTCGTTGAAAT 1614  
Db 1929 TGTGAGCAACAGAGTATGTTGATTTGAAGAGAAAGAAATCTTGAAGTTACTTGAGCT 1988  
Qy 1615 TCTGCGAAACATPAGGGATGAAGAACT---ATGGCTTTTGTGAAACTAAGAAAAGC 1671  
Db 1989 TCTAGGCCATTATCAAGAGTCAAGATCTGTCAATTATTTCTGGATAGCAGGAACATGC 2048  
Qy 1672 AGATTTTACTGCACTTTCTTTGTCAGAGAAAATATCAACTCAAGATATCCATGGTA 1731  
Db 2049 TGATGGTCTCTTAAAGATTTAATGAGAGCATCTTATCCCTTGATGTCTCTTCATGAGG 2108  
Qy 1732 TCGGGAACAGAGAGCGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAGTCCCAAGT 1791  
Db 2109 CATTTGATCAATATGACAGAGATAGCATCATTAATGACITTAAGATGGACCTGCAACT 2168  
Qy 1792 TCTTGTGTACTCTCGATAGTGCAGAGGCGGTGATATTGAAATGTGCAACATGTTAT 1851  
Db 2169 TCTTGTGGTACTCTCTGTCTGCTGCGGAGGTCTAGATGTGAACATCTGATTTCTGTAGT 2228  
Qy 1852 CAATTTTGTATCTCTCTTCTACCATTTGATGATGATGATGATGATGATGATGATG 1911  
Db 2229 AATATTATAGCTGCCCCAACCATATATGAGGATTTATGACAGAGCAGGCGGAGTGAAG 2288  
Qy 1912 TTGTGGGAATATGCGAGAGCAATTTCCPTTTT 1944  
Db 2289 AGCAGGAAACAAGGTTATGCTTATACTTTTAT 2321

RESULT 12  
US-09-543-681A-1401  
; Sequence 1401, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 1401  
; LENGTH: 1410  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-1401

Query Match 5.7%; Score 127.6; DB 4; Length 1410;  
Best Local Similarity 47.1%; Pred. No. 5.8e-27;  
Matches 514; Conservative 0; Mismatches 554; Indels 24; Gaps 3;  
Qy 880 TTTTGAAGAGCTAAATCTCTGTGACAGACACTGATATCAACATTTGCTTAAAGCTGGTTATAC 939  
Db 21 TTTTACATCGCTTGGTTTAAAGTACGAGCGCTTCTCGCGCTATTGATGAACAAGGGTATAA 80  
Qy 940 TAAGCTTACTCTGTGCAAAAATACAGATTTCTATCATCTATGACAGGACGATTTGAT 999  
Db 81 AACCCCAACCCCTATTCAACACAGCGGATTTAGCCGATTTTGGCGGTTAAGACGTATT 140  
Qy 1000 GGCTTGTGCTCAACAGGGTCTGGGAAGACTGGGGCTTTCTCTACCAATTTGGCTCA 1059  
Db 141 AGCAAGTGCAACAGGGTACGGGAAACAGGCGCGATTTACATTACCGATACCTGAAA 200  
Qy 1060 TATGATGCATGATGAATAACTGCGAGTGTCTTTTAAAGAGTTGACAGGAACCGAGTGTAT 1119  
Db 201 ACTC-----GCTACATCAGCAGAGAAAACGAAAGGCGGTAAAGCCAGTGTAAAGCGCT 251  
Qy 1120 TATTGTAGCACCAACTCGAGAAATGGTCAACAGATTTATTTTGGAGCCAGAAAATTTTC 1179

Db 252 TATTTTAAACCCCTACACGTAATTTAGTCGCAAAATTTGCCGATACATATAAAGCTTATAG 311  
Qy 1180 TTTTGGGACTTTGTGTAAGAGCTGTTGTTATATATATGGGGAAACCCAGCTGGGACATTCAT 1239  
Db 312 TCGTTATTTACCTTATTCGTTTCATTTGTTGTTTGGTGGCGTTAGTATCAATTCCTCAAT 371  
Qy 1240 TCGACAAATAGTACAAGGCTGTAAATATATATTTGTCTACTCTCTGGAAGACTGATGATAT 1299  
Db 372 GATGAACATACGAGGTGGCGTGTGATTTTGTATGCAACACCGGGCGCTTACTTGATCT 431  
Qy 1300 CATAGCAAGAAAAGATTTGGTCTCAACAGAGATCAAAATCTTTAGTTTTCGATGAAGCTGA 1359  
Db 432 TGAACATCAAAATGCTGTGCTGATTAATCTCGCGTTGAAGTATTGGTGTAGTAGAAGCTGA 491  
Qy 1360 TCCATGTTGGATATCGGTTTGGTCCAGAAATGAAGAAATTAATTTCTTCCAGGAAAT 1419  
Db 492 TCGTATGTTAGATATGGGATTTATTCATGATA-----TTCTGCGAGTCATCAA 539  
Qy 1420 GCCATCAAGGAACAGCGCCAAACCTTATGTTTCACTGCAACCTTTTCCAGAGGAAATTC 1479  
Db 540 TAAGTTACCGAAAACGACAAATTTACTCTTTTCAGGACCTTTTCAAAGAGATAAC 599  
Qy 1480 AAGGTTGGCTGACAGGTTTAAAGTCAAAATTAATCTGTTTGTGCTGTTGGACAAGTGG 1539  
Db 600 AGGCT---TGCTAACTCACTACTCAACAAATCTATAAGTATTGCTGTTGCCCCAAAA 656  
Qy 1540 TGGAGCATGTAGAGATGTTTCAGCAGACCGTTCTTCAAGTTGGCCAGTTCTCAAAAAGAG 1599  
Db 657 CTCAGCGGCTGAATCTGTTGATCAATATGTCATTTAGTGGATATAAAGGTTAAGACAGA 716  
Qy 1600 AAAGCTGTTGAAAATCTGCGAAACATPAGGGGATGAAGAACTATGCTCTTTGTTGAAAC 1659  
Db 717 GCTGTTATCACACTTAATTTGCTTAGAAAATTTGGCCTCAAGTTCTTTATTTTACTGAAAC 776  
Qy 1660 TAGAAAAAAGCAGATTTTACTGCAACTTTTCTTGTCAAGAAAAATATCAACTACAAG 1719  
Db 777 TAAACATGGTCCGAATAAATTTAGCTGAACATCTTAATTTAGATGGCATCAATCAGCGGC 836  
Qy 1720 TATCCATGGTGATCGGGAACAGAGAGAGCGGAGCAAGCTCTTGGAGATTTTCCGTTTGG 1779  
Db 837 GATCCATGGTAATAAAGCCAGAGCTAGAACACGCTGTTTACGTGATTTTAAAGACGG 896  
Qy 1780 AAGTGGCCAGTTCTTGTCTACTTCTAGTAGTGCAGAGGCTGGATTTGAAAATGT 1839  
Db 897 TAAATTAAGAGCGTTAGTTGCCACTGATATTGACGGCGGAGCTTGATATTGACCAACT 956  
Qy 1840 GCAACATGTTATCAATTTTGTGATCTTCTTACCATTGATGATATGTTTCATCGAATGG 1899  
Db 957 TCTTATGTTGGTCAACTTTTGAATTTACCGAGGTAGCAGAGATACGTTCAATCGTATTGG 1016  
Qy 1900 GCGTACTGCTGTTTGGGAATATCTGCGAGAGCAATTTCCCTTTTGTGATCTTGAATGGGA 1959  
Db 1017 TCGAACAGGGCGTGGCGGCAACAGGTAAAGCGATATCTTTAGTTTGTGTTGATGAACA 1076  
Qy 1960 TAACCATTTAGC 1971  
Db 1077 TGGCTTATTAGC 1088

RESULT 13  
US-09-557-884-1/c  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue

CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Query March 5.6%; Score 125.6; DB 4; Length 1830121;  
Best Local Similarity 47.3%; Pred. No. 1.3e-24;  
Matches 506; Conservative 0; Mismatches 534; Indels 30; Gaps 3;

881 TTTGAAGAAGCTAATCTCTGCACACACTGAATTAACCAATGCTGTAAGCTGTTACT 940  
Db 444736 TTTGAACAATTCGATCTTCTCTGAGCTTTTAAAGGCACTTGAGAAAAAGGTTATTC 444677  
QY 941 AAGCTTACTCTGTGCAAAATACAGTATTCCTATCATCTGCGAGCAGAGATTGATG 1000  
Db 444676 CGCCCAACAGCTATTCAATGGAGCCATTCCTGCCCAATGGAAGAGATGATTA 444617  
QY 1001 GCTTGCTCAACAGAGCTCTGGGAAGACTCGGCTTTCTCTACCAATTTTGGCTCAT 1060  
Db 444616 GGCTCGGACCAACCGGAACAGGAAACTCTGCTTTTATTACCTGCGCTACAACAT 444557  
QY 1061 AFGATGCTGATGATTAATCTCCAGTCTGTTTAAAGAGTTGCAGGACAGAGTGATT 1120  
Db 444556 TTATTGGATTATCCAGCGCGTAAACCGGCCCC-----ACCACGTATTTG 444512  
QY 1121 ATTGTAGCAACCACTCGAGAATTGCTCAACAGATTTTATTGGAGCCAGAAAAATTTCT 1180  
Db 444511 GTATTACACCAACCCGCTGACTGCAATGCAATGCGTGAACAGCGGAAGATTAGCG 444452  
QY 1181 TTGGGACTGTGTAAGAGCTGTTGTTATATATATGTTGCTACTCTCGAAGACTGATGATATC 1240  
Db 444451 CAGTTTCAACCAATTAATATTCGCAATTTACAGGTGCGTGGCGTATCAAAATCACCGT 444392  
QY 1241 CGAATAATAGTACAGGCTGTAATATATATGTTGCTACTCTCGAAGACTGATGATATC 1300  
Db 444391 GATGTATTCAATACCAATCAAGATTTGTTGGTGGCTACCGACCGCTTTGTTGCAATAC 444332  
QY 1301 ATAGGCAAGAAAGATTGGTCTCAACAGATCAAAATCTAGTTTGGATGAGAGCTGAT 1360  
Db 444331 ATTAAGGAAGAAAAATTTGATTCGGTTCGGTTCGAAATGCTGAATTTTGTATGAAGCGAT 444272  
QY 1361 CCAATGTTGATGATGTTGGTTCGCAATAGAGATTAATTTCTTGGCCAGGAATG 1420  
Db 444271 AGAATGTTGCAATGGATTTGGGCAAGATGCGGAAAAATTCGAGCTGAAACCGCTGG 444212  
QY 1421 CCATCAAGGAACAGCGCCAAACCCCTTATGTTTCAGTGCACATTTTCCAGAGGAATCAA 1480

## RESULT 14

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae Rd Genome, Fragments

Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

QY	1718	AGTATCCATGGTGATCGGGACACAGAGAGCGGAGCAGCTCTTCGAGATTTTCGCTTT	1777
DB	443923	TATTTAGAAAGCGCAATGGCAAACTCAACGTAAACAATGCCATTGATAATTGAATCA	443864
QY	1778	GGAAAGTCCCCAGTTCTTGTTGTCTCACTAGTCTGCCAGAGGGCTGCATATTGAAAT	1837
DB	443863	GGTATTGTGACGGTATTGGTTGCAACAGATGTGGCTGCACGTGGTATTGATATTGACGAT	443804
QY	1838	GTGCAACATGTTATCAATTTTGTCTTCCTCTTACCATTTGATGAATATGTTTCATCGAAAT	1897
DB	443803	GTAAGCCACGTGATGAATTTTGTATTCCTCTATGTGCGGATTAATTTTTCATTCGAAAT	443744
QY	1898	GGCGCTACTGTCGTGTGGGAATCTGGCAGACCAATTTCCCTTTTTCGA	1947
DB	443743	GGACGTAACGGCGAGCTTGGCAAAAAGGACACGGCAGTCTCTTTTCTCGA	443694

RESULT 15  
PCT-US96-05320A-894  
Sequence 894, Application PC/TUS9605320A  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences  
APPLICANT: 9410 Key West Avenue  
APPLICANT: Rockville, MD 20850  
APPLICANT: United States of America  
APPLICANT: Johns Hopkins University  
APPLICANT: 720 Rutland Avenue  
APPLICANT: Baltimore, MD 21205  
APPLICANT: United States of America  
APPLICANT: Mark D. Adams  
APPLICANT: Owen White  
APPLICANT: Hamilton O. Smith  
APPLICANT: J. Craig Venter  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20003-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/05320A  
FILING DATE: April 22, 1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: June 7, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Eric K. Steffe  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.014PC01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 894:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1254 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
PCT-US96-05320A-894



Query Match				5.6%;	Score 123.6;	DB 5;	Length 1254;
Best Local Similarity				47.1%;	Pred. No. 7.9e-26;		
Matches 495;				Conservative 0;	Mismatches 539;	Indels 18;	Gaps 3;
Qy	921	TTGCTAAAGCTGGTTTACTAAGCTTACTCTCTGTCGCAAAATACAGTATTCCTATCATAC	980				
Db	80	TGGCAAAAAGGCTTTGATTTTGTACCCCAATTCAGGCTTTATCCTGCTCATCAGTT	139				
Qy	981	TTGCAGACGAGATTGATGGCTTGTCTCAACAGGGTCTGGAGACTGGCGCTTTTC	1040				
Db	140	TAAATGGACGAGATGTCGACGACAAAGCTCAACTGTGTACAGGCAAGCAATGGCTTTT	199				
Qy	1041	TCCTACCAATTTTGGCTCATATGATGATGATGATGAATAACTGCCAGTCGTTTTAAAGAGT	1100				
Db	200	TAA CGGCTACTTTTCCACCATCTTTTAACTCACCAAGA-----TCCTAACTTAAAT	250				
Qy	1101	TGCAGAACACAGAGTATATTTAGTGTAGCACCAACTCGAGAAATGGTCAACAGATTTAT	1160				
Db	251	ATCCTCACCCCAAGAGCTTTGATTTTGTAGCCTACTCGAGAAATAGCGGTACAGATAGTA	310				
Qy	1161	TGGAAGCCAGAAAATTTCTTTTGGGACTTGTGTAGAGGCTGTGTATATATATATATATATAT	1220				
Db	311	ATGACGAGATTTCTTTCARAAAGCGAGTGGATTAAGACCGCACTTGCCTATGGTGGCG	370				
Qy	1221	CCCAGCTGGGACATTCATATTCGACAAATAGTACAAAGGCTGTATATATATATATATAT	1280				
Db	371	ATGGTTATGATAAACAACACTACAGCGATTCAGCGTGGCGTCGATATTTTGTATGGTACGA	430				
Qy	1281	CTGGAAGACTGATCGATATCATAGGCAAGAAAGATTTGTTCTCAACAGATCAATATCT	1340				
Db	431	CGGGCGGAGTCATTTGATTTGTGAACAGCGGTAAATGGTTAGATGAATCCAGTTG	490				
Qy	1341	TAGTTTTGGATGAAGCTGATCGCATGTGTGATATGGTTTTGGTCCAGAAATGAAGAAGT	1400				
Db	491	TCGTGTAGATGAAGCAGATCGAATGTTGATCTTGGGTTTATCCGTGATATTCGTTATT	550				
Qy	1401	TAAATTTCTTGGCCAGGATGCATCAAGGACAGCGCCAAACCTTATGTTCAAGTGCAA	1460				
Db	551	TATTGCGTAA-----ATGCCCGCTCCGCAAGCTCGTTTAAACGATGTTATTTTCAGCGA	604				
Qy	1461	CTTTTCCAGAGAAATTCARAGTTGGCTGCAGAGTTTTTAAAGTCAAATATCTGTTTG	1520				
Db	605	CGCTTTCTTATAAGTGGTGAATTAGCATTTGAAGATATGAATGAACCTGA--ATATA	661				
Qy	1521	TTGCTGTGGACAGTGGGTGGAGCAGTAGAGATGTTACAGCAGACCGTTCTCCAGTTG	1580				
Db	662	TTGAATTTGAACCCAGAACAAAAACCGGGCACCGCAATTAAGAAGAACTTTTTTATCCAT	721				
Qy	1581	GCCAGTTCTCAAAAGAGAAAGCTCGTTGAAATTTCTCGAAACATAGGGGATGAAGAA	1640				
Db	722	CTAATCAGGATAAATGSCACTTCTCTTAACCTTAAATGGAGATGAATGSCCTGAACGCT	781				
Qy	1641	CTATGGTCTTTGTTGAAACTAAGAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAG	1700				
Db	782	GTATTTGATTTGCGAATACGAAACATCGTTGTGAAGAAATTTGGGGCTATTTGGCGGCTG	841				
Qy	1701	AAAAAATATCACTACAGTATCCATGTTGATCGGGAACAGAGAGCGGGAGCAAGCTC	1760				
Db	842	ATGGGCATCGGTGCGTTTACTGACTGCGGATGTAGCACGAAAAACGTTTATCGTTAT	901				
Qy	1761	TTGAGATTTTTCGTTTGGAAAGTCCCGAGTTCTTGTGTACTTTCAGTAGCTGCCAGAG	1820				
Db	902	TAAAAAATTTACTGATGGTGTATTTGGGATAATTTAGTGGCAACAGATGTGCTGCTG	961				
Qy	1821	GGCTGGATTTGAAATGTGCAACATGTTATCAATTTTGTATCTCTTCTTACCATTGATG	1880				
Db	962	GCTTGCAATTTCTGATGTGACGATGTTTCAATATATGATTTACCCGATGATCGGAAG	1021				
Qy	1881	AATATGTTTCAGATTTGGGCTACTGCTCGTTGTGGGAATCTATGGCAGAGCAATTTTCCT	1940				
Db	1022	ATTATGTTTCCAGGATTTGGGCTACTGACCGAGCGGGAAAGTGGTGTTCGATTAGTT	1081				
Qy	1941	TTTTTGATCTTGAATCGGATAAACCAATTTAGCA	1972				

Db 1082 TCGCTTGTGAAGATATCGGATGAATTTACCA 1113

Search completed: May 9, 2004, 16:55:37

Job time : 194.964 secs